



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Chandrasegaran, Srinivasan
- (ii) TITLE OF INVENTION: A GENERAL METHOD TO CLONE HYBRID RESTRICTION ENDONUCLEASES USING lig GENE
- (iii) NUMBER OF SEQUENCES: 35
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: Cushman Darby & Cushman L.L.P.  
(B) STREET: 1100 New York Avenue, NW, Ninth Floor, East Tower  
(C) CITY: Washington  
(D) STATE: DC  
(E) COUNTRY: USA  
(F) ZIP: 20005-3918
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: US  
(B) FILING DATE: 20-DEC-1995  
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Kokulis, Paul N.  
(B) REGISTRATION NUMBER: 16,773  
(C) REFERENCE/DOCKET NUMBER: PNK/4130/213779/DJP
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 202-861-3000  
(B) TELEFAX: 202-822-0944  
(C) TELEX: 6714627 CUSH

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

54 53

Leu Tyr Xaa Cys Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa  
1 5 10 15

Xaa Xaa His Xaa Xaa Xaa His  
20

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

*a a /*  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Leu Tyr Xaa Cys Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa  
1 5 10 15

Xaa Xaa His Xaa Xaa Xaa Xaa His  
20

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 25 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

*a a /*  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Leu Tyr Xaa Cys Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa  
1 5 10 15

Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His  
20 25

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

*54*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Xaa Tyr Xaa Cys Xaa Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa  
1 5 10 15

Leu Xaa Xaa His Xaa Xaa Xaa His  
20

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 25 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

*AA*  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Xaa Tyr Xaa Cys Xaa Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa  
1 5 10 15

Leu Xaa Xaa His Xaa Xaa Xaa Xaa His  
20 25

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 26 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Xaa Tyr Xaa Cys Xaa Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa  
1 5 10 15

Leu Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His  
20 25

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 25 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

*54 55*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Xaa Tyr Xaa Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa  
Xaa 1 5 10 15

Xaa Leu Xaa Xaa His Xaa Xaa Xaa His  
20 25

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 26 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

*Ra (*  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Xaa Tyr Xaa Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa  
Xaa 1 5 10 15

Xaa Leu Xaa Xaa His Xaa Xaa Xaa Xaa His  
20 25

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 27 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Xaa Tyr Xaa Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa  
Xaa 1 5 10 15

Xaa Leu Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His  
20 25

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 amino acids
- 51 SK*

- (B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Phe Xaa Cys Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa  
Leu 1 5 10 15

Xaa Xaa His Xaa Xaa Xaa His  
20

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Phe Xaa Cys Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa  
Leu 1 5 10 15

Xaa Xaa His Xaa Xaa Xaa Xaa His  
20

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 25 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Phe Xaa Cys Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa  
Leu 1 5 10 15

Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His  
20 25

(2) INFORMATION FOR SEQ ID NO:13:

58 57

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Xaa Phe Xaa Cys Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa  
1 5 10 15

Leu Xaa Xaa His Xaa Xaa Xaa His  
20

(2) INFORMATION FOR SEQ ID NO:14:

- 19 (*
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 25 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Xaa Phe Xaa Cys Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa  
1 5 10 15

Leu Xaa Xaa His Xaa Xaa Xaa Xaa His  
20 25

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 26 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Xaa Phe Xaa Cys Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa  
1 5 10 15

Leu Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His  
20 25

*59 58*

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 25 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Phe Xaa Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa  
Xaa 1 5 10 15

Xaa Leu Xaa Xaa His Xaa Xaa Xaa His  
20 25

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 26 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Phe Xaa Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa  
Xaa 1 5 10 15

Xaa Leu Xaa Xaa His Xaa Xaa Xaa Xaa His  
20 25

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 27 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Phe Xaa Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Phe Xaa Xaa Xaa  
Xaa 1 5 10 15

60 59

Xaa Leu Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His  
20 25

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

NNNNNNNNNN NNNNNNNNNN NNNNNGAGGA GGCTNNNNN  
39

(2) INFORMATION FOR SEQ ID NO:20:

- 61*
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

NNNNNNNNNN NNNNNNNNNN NNNNNGAGGG ATGTNNNNN  
39

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 74 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GAGCTCCTTT GGATCCAAGC TTCCCGGGGA AGAATTGAG GAGGCTCTCG AGGTGACTT  
60

CCTCTAGAGG TACC  
74

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
- U1 60*

- (A) LENGTH: 74 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GAGCTCCTTT GGATCCAAGC TTCCCGGGGA AGAATTGAG GAGGCTCTCG AGGTCGACTT  
60

CCTCTAGAGG TACC  
74

(2) INFORMATION FOR SEQ ID NO:23:

- AA*
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 74 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CUCGACCAAA CCUAGGUUCG AAGGGCCCCU UCUUAAGCUC CUCCGAGAGC UCCAGCUGAA  
60

GGAGAUCUCC AUGG  
74

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 87 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGATCCAAGC TTAGCGATCT GCCTGCAGGT CGACTCTAGC CAGGGGCGGG GTGGTCTAGT  
60

ATCGTTCAAT GATACTTCAT GGAATTG  
87

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:

*62 L1*

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CCCCCTGAAGG AGATATAACAT ATG  
23

(2) INFORMATION FOR SEQ ID NO:26:

- Ra'
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 25 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GGACTAGTCC CTTCTTATTTC TGGTG  
25

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 15 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Ser  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 14 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

63/2

CTGACGGGGG CCAA

14

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CTAGTTGGCC CCCGT

15

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CTAGACGGGG GAGGCAGCAG TCAA

24

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CTAGTTGACT GCCGCCTCCC CCGT

24

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

64 63

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Gly Gly Gly Gly Ser  
1 5

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CTAGACGGGG GAGGCAGGCAAG TGGAGGTGGC GGATCACAA  
39

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CTAGTTGTGA TCCGCCACCT CCACCGTCGC CTCCCCCGT  
39

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser  
1 5 10

65

64

08/575, 361

**NOTICE TO COMPLY WITH REQUIREMENTS FOR PATENT APPLICATIONS CONTAINING  
NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURES**

The nucleotide and/or amino acid sequence disclosure contained in this application does not comply with the requirements for such a disclosure as set forth in 37 CFR 1.821 - 1.825 for the following reason(s):

- 1. This application clearly fails to comply with the requirements of 37 CFR 1.821 - 1.825. Applicant's attention is directed to these regulations, published at 1114 OG 29, May 15, 1990 and at 55 FR 18230, May 1, 1990.
- 2. This application does not contain, as a separate part of the disclosure on paper copy, a "Sequence Listing" as required by 37 CFR 1.821(c).
- 3. A copy of the "Sequence Listing" in computer readable form has not been submitted as required by 37 CFR 1.821(e).
- 4. A copy of the "Sequence Listing" in computer readable form has been submitted. However, the content of the computer readable form does not comply with the requirements of 37 CFR 1.822 and/or 1.823, as indicated on the attached marked-up copy of the "Raw Sequence Listing."
- 5. The computer readable form that has been filed with this application has been found to be damaged and/or unreadable as indicated on the attached CRF Diskette Problem Report. A substitute computer readable form must be submitted as required by 37 CFR 1.825(d).
- 6. The paper copy of the "Sequence Listing" is not the same as the computer readable form of the "Sequence Listing" as required by 37 CFR 1.821(e).
- 7. Other: \_\_\_\_\_

**Applicant must provide:**

- An initial or substitute computer readable form (CRF) copy of the "Sequence Listing"
- An initial or substitute paper copy of the "Sequence Listing", as well as an amendment directing its entry into the specification
- A statement that the content of the paper and computer readable copies are the same and, where applicable, include no new matter, as required by 37 CFR 1.821(e) or 1.821(f) or 1.821(g) or 1.825(b) or 1.825(d)

For questions regarding compliance with these requirements, please contact:

For Rules Interpretation, call (703) 308-1123

For CRF submission help, call (703) 308-4212

For PatentIn software help, call (703) 308-6856

Please return a copy of this notice with your response.